#2

OIPE

ENTERED

```
 3 <110> APPLICANT: Shimkets, Richard
 4         Fernandes, Elma
 5         Boldog, Ferenc
 7 <120> TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
 9 <130> FILE REFERENCE: 15966-551
11 <140> CURRENT APPLICATION NUMBER: 09/998,966
12 <141> CURRENT FILING DATE: 2001-10-31
14 <150> PRIOR APPLICATION NUMBER: 09/569,269
15 <151> PRIOR FILING DATE: 2000-05-11
17 <150> PRIOR APPLICATION NUMBER: 60/134,315
18 <151> PRIOR FILING DATE: 1999-05-14
20 <150> PRIOR APPLICATION NUMBER: 60/175,744
21 <151> PRIOR FILING DATE: 2000-01-12
23 <150> PRIOR APPLICATION NUMBER: 60/188,274
24 <151> PRIOR FILING DATE: 2000-03-10
26 <160> NUMBER OF SEQ ID NOS: 52
28 <170> SOFTWARE: PatentIn Ver. 2.0
30 <210> SEQ ID NO: 1
31 <211> LENGTH: 670
32 <212> TYPE: DNA
33 <213> ORGANISM: Homo sapiens
35 <220> FEATURE:
36 <221> NAME/KEY: CDS
37 <222> LOCATION: (130)..(639)
39 <400> SEQUENCE: 1
40 ccattggccg gcgtccccgc cccagcgaac ccggccccgc ccccgaggcg ccccattggc 60
42 cccgccgcgc gaaggcagag ccgcggacgc ccgggagcga cgagcgcgca gcgaaccggg 120
44 tgccgggtc atg cgc cgc cgc ctg tgg ctg ggc ctg gcc tgg ctg ctg ctg 171
45           Met Arg Arg Arg Leu Trp Leu Gly Leu Ala Trp Leu Leu Leu
46             1               5                   10
48 gcg cgg gcg ccg gac gcc gcg gga acc ccg agc gcg tcg cgg gga ccg 219
49 Ala Arg Ala Pro Asp Ala Ala Gly Thr Pro Ser Ala Ser Arg Gly Pro
50  15                  20                  25                  30
52 cgc agc tac ccg cac ctg gag ggc gac gtg cgc tgg cgg cgc ctc ttc 267
53 Arg Ser Tyr Pro His Leu Glu Gly Asp Val Arg Trp Arg Arg Leu Phe
54                 35                  40                  45
56 tcc tcc act cac ttc ttc ctg cgc gtg gat ccc ggc ggc cgc gtg cag 315
57 Ser Ser Thr His Phe Phe Leu Arg Val Asp Pro Gly Gly Arg Val Gln
58             50                  55                  60
60 ggc acc cgc tgg cgc cac ggc cag gac agc atc ctg gag atc cgc tct 363
61 Gly Thr Arg Trp Arg His Gly Gln Asp Ser Ile Leu Glu Ile Arg Ser
62          65                  70                  75
64 gta cac gtg ggc gtc gtg gtc atc aaa gca gtg tcc tca ggc ttc tac 411
```

```
65 Val His Val Gly Val Val Val Ile Lys Ala Val Ser Ser Gly Phe Tyr
66     80                  85                  90
68 gtg gcc atg aac cgc cgg ggc cgc ctc tac ggg tcg cga ctc tac acc    459
69 Val Ala Met Asn Arg Arg Gly Arg Leu Tyr Gly Ser Arg Leu Tyr Thr
70 95                  100                  105                  110
72 gtg gac tgc agg ttc cgg gag cgc atc gaa gag aac ggc cac aac acc    507
73 Val Asp Cys Arg Phe Arg Glu Arg Ile Glu Glu Asn Gly His Asn Thr
74               115                  120                  125
76 tac gcc tca cag cgc tgg cgc cgc cgc ggc cag ccc atg ttc ctg gcg    555
77 Tyr Ala Ser Gln Arg Trp Arg Arg Arg Gly Gln Pro Met Phe Leu Ala
78               130                  135                  140
80 ctg gac agg agg ggg ggg ccc cgg cca ggc ggc cgg acg cgg cgg tac    603
81 Leu Asp Arg Arg Gly Gly Pro Arg Pro Gly Gly Arg Thr Arg Arg Tyr
82               145                  150                  155
84 cac ctg tcc gcc cac ttc ctg ccc gtc ctg gtc tcc tgaggccctg         649
85 His Leu Ser Ala His Phe Leu Pro Val Leu Val Ser
86     160                  165                  170
88 agaggccggc ggctccccaa g                                            670
91 <210> SEQ ID NO: 2
92 <211> LENGTH: 170
93 <212> TYPE: PRT
94 <213> ORGANISM: Homo sapiens
96 <400> SEQUENCE: 2
97 Met Arg Arg Arg Leu Trp Leu Gly Leu Ala Trp Leu Leu Leu Ala Arg
98  1                  5                  10                  15
100 Ala Pro Asp Ala Ala Gly Thr Pro Ser Ala Ser Arg Gly Pro Arg Ser
101             20                  25                  30
103 Tyr Pro His Leu Glu Gly Asp Val Arg Trp Arg Arg Leu Phe Ser Ser
104             35                  40                  45
106 Thr His Phe Phe Leu Arg Val Asp Pro Gly Gly Arg Val Gln Gly Thr
107     50                  55                  60
109 Arg Trp Arg His Gly Gln Asp Ser Ile Leu Glu Ile Arg Ser Val His
110 65                  70                  75                  80
112 Val Gly Val Val Val Ile Lys Ala Val Ser Ser Gly Phe Tyr Val Ala
113             85                  90                  95
115 Met Asn Arg Arg Gly Arg Leu Tyr Gly Ser Arg Leu Tyr Thr Val Asp
116             100                  105                  110
118 Cys Arg Phe Arg Glu Arg Ile Glu Glu Asn Gly His Asn Thr Tyr Ala
119             115                  120                  125
121 Ser Gln Arg Trp Arg Arg Arg Gly Gln Pro Met Phe Leu Ala Leu Asp
122     130                  135                  140
124 Arg Arg Gly Gly Pro Arg Pro Gly Gly Arg Thr Arg Arg Tyr His Leu
125 145                  150                  155                  160
127 Ser Ala His Phe Leu Pro Val Leu Val Ser
128             165                  170
131 <210> SEQ ID NO: 3
132 <211> LENGTH: 1680
133 <212> TYPE: DNA
134 <213> ORGANISM: Homo sapiens
```

RAW SEQUENCE LISTING   DATE: 01/29/2002
PATENT APPLICATION: US/09/998,966   TIME: 15:45:55

Input Set : N:\Crf3\RULE60\09998966.txt
Output Set: N:\CRF3\01292002\I998966.raw

```
136 <220> FEATURE:
137 <221> NAME/KEY: CDS
138 <222> LOCATION: (177)..(1655)
140 <400> SEQUENCE: 3
141 tgacaggccg gccggtgagg cgccgccggg agaggccgcg acggagctcc cagaccggcc 60
143 atgggctgag acacgtcctc gccgagcagt gacccttccg taccccacca gaacatgccc 120
145 gggtgacctc ctcccagatc ttccttgtgg ccttcctcgc ccactccagt gacact atg 179
146                                                                 Met
147                                                                   1
149 cac ccc cac cgt gac ccg aga ggc ctc tgg ctc ctg ctg ccg tcc ttg 227
150 His Pro His Arg Asp Pro Arg Gly Leu Trp Leu Leu Leu Pro Ser Leu
151             5                   10                  15
153 tcc ctg ctg ctt ttt gag gtg gcc aga gct ggc cga gcc gtg gtt agc 275
154 Ser Leu Leu Leu Phe Glu Val Ala Arg Ala Gly Arg Ala Val Val Ser
155             20                  25                  30
157 tgt cct gcc gcc tgc ttg tgc gcc agc aac atc ctc agc tgc tcc aag 323
158 Cys Pro Ala Ala Cys Leu Cys Ala Ser Asn Ile Leu Ser Cys Ser Lys
159     35                  40                  45
161 cag cag ctg ccc aat gtg ccc cat tcc ttg ccc agt tac aca gca cta 371
162 Gln Gln Leu Pro Asn Val Pro His Ser Leu Pro Ser Tyr Thr Ala Leu
163 50                  55                  60                  65
165 ctg gac ctc agt cac aac aac ctg agc cgc ctg cgg gcc gag tgg acc 419
166 Leu Asp Leu Ser His Asn Asn Leu Ser Arg Leu Arg Ala Glu Trp Thr
167                 70                  75                  80
169 ccc acg cgc ctg acc caa ctg cac tcc ctg ctg ctg agc cac aac cac 467
170 Pro Thr Arg Leu Thr Gln Leu His Ser Leu Leu Leu Ser His Asn His
171                 85                  90                  95
173 ctg aac ttc atc tcc tct gag gcc ttt tcc ccg gta ccc aac ctg cgc 515
174 Leu Asn Phe Ile Ser Ser Glu Ala Phe Ser Pro Val Pro Asn Leu Arg
175             100                 105                 110
177 tac ctg gac ctc tcc tcc aac cag ctg cgt aca ctg gat gag ttc ctg 563
178 Tyr Leu Asp Leu Ser Ser Asn Gln Leu Arg Thr Leu Asp Glu Phe Leu
179     115                 120                 125
181 ttc agt gac ctg caa gta ctg gag gtg ctg ctg ctc tac aat aac cac 611
182 Phe Ser Asp Leu Gln Val Leu Glu Val Leu Leu Leu Tyr Asn Asn His
183 130                 135                 140                 145
185 atc atg gcg gtg gac cgg tgc gcc ttc gat gac atg gcc cag ctg cag 659
186 Ile Met Ala Val Asp Arg Cys Ala Phe Asp Asp Met Ala Gln Leu Gln
187                 150                 155                 160
189 aaa ctc tac ttg agc cag aac cag atc tct cgc ttc cct ctg gaa ctg 707
190 Lys Leu Tyr Leu Ser Gln Asn Gln Ile Ser Arg Phe Pro Leu Glu Leu
191             165                 170                 175
193 gtc aag gaa gga gcc aag cta ccc aaa cta acg ctc ctg gat ctc tct 755
194 Val Lys Glu Gly Ala Lys Leu Pro Lys Leu Thr Leu Leu Asp Leu Ser
195             180                 185                 190
197 tct aac aag ctg aag aac ttg cca ttg cct gac ctg cag aag ctg ccg 803
198 Ser Asn Lys Leu Lys Asn Leu Pro Leu Pro Asp Leu Gln Lys Leu Pro
199     195                 200                 205
201 gcc tgg atc aag aat ggg ctg tac cta cat aac aac ccc ctg aac tgc 851
```

RAW SEQUENCE LISTING   DATE: 01/29/2002
PATENT APPLICATION: US/09/998,966   TIME: 15:45:55

Input Set : N:\Crf3\RULE60\09998966.txt
Output Set: N:\CRF3\01292002\I998966.raw

```
202 Ala Trp Ile Lys Asn Gly Leu Tyr Leu His Asn Asn Pro Leu Asn Cys
203 210             215             220             225
205 gac tgt gag ctc tac cag ctg ttt tca cac tgg cag tat cgg cag ctg      899
206 Asp Cys Glu Leu Tyr Gln Leu Phe Ser His Trp Gln Tyr Arg Gln Leu
207             230             235             240
209 agc tcc gtg atg gac ttt caa gag gat ctg tac tgc atg aac tcc aag      947
210 Ser Ser Val Met Asp Phe Gln Glu Asp Leu Tyr Cys Met Asn Ser Lys
211             245             250             255
213 aag ctg cac aat gtc ttc aac ctg agt ttc ctc aac tgt ggc gag tac      995
214 Lys Leu His Asn Val Phe Asn Leu Ser Phe Leu Asn Cys Gly Glu Tyr
215         260             265             270
217 aag gag cgt gcc tgg gag gcc cac ctg ggt gac acc ttg atc atc aag     1043
218 Lys Glu Arg Ala Trp Glu Ala His Leu Gly Asp Thr Leu Ile Ile Lys
219     275             280             285
221 tgt gac acc aag cag caa ggg atg acc aag gtg tgg gtg aca cca agt     1091
222 Cys Asp Thr Lys Gln Gln Gly Met Thr Lys Val Trp Val Thr Pro Ser
223 290             295             300             305
225 aat gaa cgg gtg cta gat gag gtg acc aat ggc aca gtg agt gtg tct     1139
226 Asn Glu Arg Val Leu Asp Glu Val Thr Asn Gly Thr Val Ser Val Ser
227             310             315             320
229 aag gat ggc agt ctt ctt ttc cag cag gtg cag gtc gag gac ggt ggt     1187
230 Lys Asp Gly Ser Leu Leu Phe Gln Gln Val Gln Val Glu Asp Gly Gly
231             325             330             335
233 gtg tat acc tgc tat gcc atg gga gag act ttc aat gag aca ctg tct     1235
234 Val Tyr Thr Cys Tyr Ala Met Gly Glu Thr Phe Asn Glu Thr Leu Ser
235         340             345             350
237 gtg gaa ttg aaa gtg cac aat ttc acc ttg cac gga cac cat gac acc     1283
238 Val Glu Leu Lys Val His Asn Phe Thr Leu His Gly His His Asp Thr
239     355             360             365
241 ctc aac aca gcc tat acc acc cta gtg ggc tgt atc ctt agt gtg gtc     1331
242 Leu Asn Thr Ala Tyr Thr Thr Leu Val Gly Cys Ile Leu Ser Val Val
243 370             375             380             385
245 ctg gtc ctc ata tac cta tac ctc acc cct tgc cgc tgc tgg tgc cgg     1379
246 Leu Val Leu Ile Tyr Leu Tyr Leu Thr Pro Cys Arg Cys Trp Cys Arg
247             390             395             400
249 ggt gta gag aag cct tcc agc cat caa gga gac agc ctc agc tct tcc     1427
250 Gly Val Glu Lys Pro Ser Ser His Gln Gly Asp Ser Leu Ser Ser Ser
251             405             410             415
253 atg ctt agt acc aca ccc aac cat gat cct atg gct ggt ggg gac aaa     1475
254 Met Leu Ser Thr Thr Pro Asn His Asp Pro Met Ala Gly Gly Asp Lys
255         420             425             430
257 gat gat ggt ttt gac cgg cgg gtg gct ttc ctg gaa cct gct gga cct     1523
258 Asp Asp Gly Phe Asp Arg Arg Val Ala Phe Leu Glu Pro Ala Gly Pro
259     435             440             445
261 ggg cag ggt caa aac ggc aag ctc aag cca ggc aac acc ctg cca gtg     1571
262 Gly Gln Gly Gln Asn Gly Lys Leu Lys Pro Gly Asn Thr Leu Pro Val
263 450             455             460             465
265 cct gag gcc aca ggc aag ggc caa cgg agg atg tcg gat cca gaa tca     1619
266 Pro Glu Ala Thr Gly Lys Gly Gln Arg Arg Met Ser Asp Pro Glu Ser
```

RAW SEQUENCE LISTING   DATE: 01/29/2002
PATENT APPLICATION: US/09/998,966   TIME: 15:45:55

Input Set : N:\Crf3\RULE60\09998966.txt
Output Set: N:\CRF3\01292002\I998966.raw

```
267                  470                 475                 480
269 gtc agc tcg gtc ttc tct gat acg ccc att gtg gtg tgagcaggat        1665
270 Val Ser Ser Val Phe Ser Asp Thr Pro Ile Val Val
271                  485                 490
273 gggttggtgg ggaga                                                   1680
276 <210> SEQ ID NO: 4
277 <211> LENGTH: 493
278 <212> TYPE: PRT
279 <213> ORGANISM: Homo sapiens
281 <400> SEQUENCE: 4
282 Met His Pro His Arg Asp Pro Arg Gly Leu Trp Leu Leu Leu Pro Ser
283   1               5                  10                  15
285 Leu Ser Leu Leu Leu Phe Glu Val Ala Arg Ala Gly Arg Ala Val Val
286              20                  25                  30
288 Ser Cys Pro Ala Ala Cys Leu Cys Ala Ser Asn Ile Leu Ser Cys Ser
289          35                  40                  45
291 Lys Gln Gln Leu Pro Asn Val Pro His Ser Leu Pro Ser Tyr Thr Ala
292    50                  55                  60
294 Leu Leu Asp Leu Ser His Asn Asn Leu Ser Arg Leu Arg Ala Glu Trp
295  65                  70                  75                  80
297 Thr Pro Thr Arg Leu Thr Gln Leu His Ser Leu Leu Leu Ser His Asn
298              85                  90                  95
300 His Leu Asn Phe Ile Ser Ser Glu Ala Phe Ser Pro Val Pro Asn Leu
301             100                 105                 110
303 Arg Tyr Leu Asp Leu Ser Ser Asn Gln Leu Arg Thr Leu Asp Glu Phe
304         115                 120                 125
306 Leu Phe Ser Asp Leu Gln Val Leu Glu Val Leu Leu Leu Tyr Asn Asn
307     130                 135                 140
309 His Ile Met Ala Val Asp Arg Cys Ala Phe Asp Asp Met Ala Gln Leu
310 145                 150                 155                 160
312 Gln Lys Leu Tyr Leu Ser Gln Asn Gln Ile Ser Arg Phe Pro Leu Glu
313             165                 170                 175
315 Leu Val Lys Glu Gly Ala Lys Leu Pro Lys Leu Thr Leu Leu Asp Leu
316             180                 185                 190
318 Ser Ser Asn Lys Leu Lys Asn Leu Pro Leu Pro Asp Leu Gln Lys Leu
319         195                 200                 205
321 Pro Ala Trp Ile Lys Asn Gly Leu Tyr Leu His Asn Asn Pro Leu Asn
322     210                 215                 220
324 Cys Asp Cys Glu Leu Tyr Gln Leu Phe Ser His Trp Gln Tyr Arg Gln
325 225                 230                 235                 240
327 Leu Ser Ser Val Met Asp Phe Gln Glu Asp Leu Tyr Cys Met Asn Ser
328             245                 250                 255
330 Lys Lys Leu His Asn Val Phe Asn Leu Ser Phe Leu Asn Cys Gly Glu
331         260                 265                 270
333 Tyr Lys Glu Arg Ala Trp Glu Ala His Leu Gly Asp Thr Leu Ile Ile
334         275                 280                 285
336 Lys Cys Asp Thr Lys Gln Gln Gly Met Thr Lys Val Trp Val Thr Pro
337     290                 295                 300
339 Ser Asn Glu Arg Val Leu Asp Glu Val Thr Asn Gly Thr Val Ser Val
```

Input Set : N:\Crf3\RULE60\09998966.txt
Output Set: N:\CRF3\01292002\I998966.raw

```
L:637 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:962 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:963 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:986 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:987 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:1064 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:1082 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:1159 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1160 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1183 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1184 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1356 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:1374 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:1996 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (25) SEQUENCE:
```